# Evaluating Deterrence Measures in Adversary Modeling

**Carol A. Meyers**
(925) 422-1252
meyers14@llnl.gov

Deterrence is a state of mind that is best defined as "the prevention of action by fear of the consequences." This state is typically brought about by the existence of a credible threat of unacceptable counteraction. The phenomenon of deterrence is not often addressed in studies of adversary modeling, primarily because it can be difficult to quantify its effects. This is at least in part because historically, deterrence has been studied within three broad areas (see table), each of which has developed a different motivation and methodology for capturing its results.

Our work integrates across the different areas in which deterrence has been studied to present a unified framework of methods for quantitatively evaluating the effects of deterrence. The end result is a toolbox of three different approaches, which we have implemented within the Modeling the Adversary for Responsive Strategy (MARS) project, an existing LLNL adversary modeling effort.

## Project Goals

This project provides a toolbox of methods for quantifying deterrence. Specifically, we focus on three different kinds of deterrent effects: 1) actions that cause the adversary to shift to a different attack; 2) actions that cause the adversary to shift to not attacking; and 3) actions that cause an increased probability of interdiction of the adversary during an attack. Each of these topics corresponds to one area of the toolbox. Ultimately, we intend for these toolbox methods to be used in future studies of adversary modeling.

## Relevance to LLNL Mission

In recent years, LLNL has invested effort in developing expertise in methods of adversary modeling, including probabilistic risk analysis, agent-based modeling, social networks, and Bayesian inference techniques. This work extends capabilities in these areas and increases proficiency for other applications, such as the assessment of critical infrastructure. It aligns with the adversary modeling roadmap within the Engineering Systems for Knowledge and Inference (ESKI) focus area and the Threat Prevention and Response Technologies theme in the LLNL Science &Technology plan.

## FY2008 Accomplishments and Results

The first area of the toolbox concerns actions that cause the adversary to shift to a different attack. For this topic, we synthesized a decision-making model from the literature and built it into the MARS effort. The cornerstone of this model is a parameter, β, that shifts between the extremes of the adversary choosing randomly and always choosing according to its maximum multi-attribute utility. Running this model, we discov-

Areas in which deterrence has been studied.

| Area | Timeframe | Goal of deterrence |
|---|---|---|
| Crime fighting | 1800s to present | Manipulation of actors through harsh punishments |
| Negotiations between hostile nations | 1940s to present (peak during Cold War) | Manipulation of actors through threats |
| Counterterrorism | 1970s to present (peak after 9/11) | Manipulation of actors through obstacles to action |

ered a counterintuitive finding, which is that occasionally countermeasures can cause more damage than they prevent. This happens when a countermeasure is placed on a less damaging target, and the adversary responds by shifting to a more damaging target (Fig. 1).

The second toolbox area addresses actions that cause the adversary to shift to not attacking. One difficulty we found with this topic is that it can be very hard to quantify the utility of not attacking. We addressed this issue by adding a non-attack option to MARS, and varying the utility associated with that option via sensitivity analyses. We noted that the addition of non-attacks can cause certain countermeasures to perform better, which happens when the addition of the countermeasure significantly decreases the overall probability of attack.

The third toolbox area deals with actions that cause an increased probability of interdiction of the adversary. Here we used an existing agent-based model of an adversary attack on a subway station, created as part of the LLNL Vulnerability Reduction effort. In this model, the adversary enters a station and attempts to detonate an explosive charge, while patrol units simultaneously attempt to interdict the adversary (Fig. 2). We showed how the agent-based model can be used in conjunction with MARS to generate a quantitative assessment of conops countermeasures, such as the number of patrols posted on each platform and whether security cameras are used.

## Related References

1. Anthony, R., “A Calibrated Model of the Psychology of Deterrence,” *Bulletin on Narcotics*, **56**, pp. 49–64, 2004.
2. Ehrlich, I., “The Deterrent Effect of Capital Punishment: A Question of Life and Death,” *The American Economic Review*, **65**, pp. 397–417, 1975.
3. Huth, P., and B. Russett, “Deterrence Failure and Crisis Escalation,” *International Studies Quarterly*, **32**, pp. 29–45, 1988.
4. Jacobson, S., T. Karnani, and J. Kobza, “Assessing the Impact of Deterrence on Aviation Checked Baggage Screening Strategies,” *International Journal of Risk Assessment and Management*, **5**, pp. 1–15, 2005.

## FY2009 Proposed Work

Our goal is to expand upon all of the areas in the toolbox, particularly the agent-based modeling component. We plan on adding Bayesian learning techniques to the patrol decision-making processes, as well as using such models to evaluate the deterrent effects of decoy countermeasures.

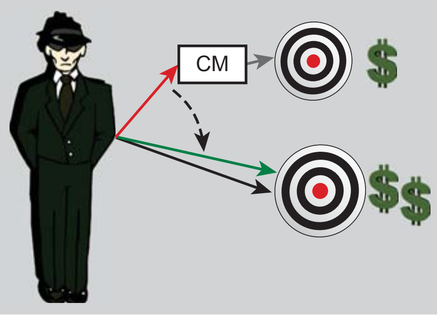

**Figure 1.** Illustration of model result: when a countermeasure is placed on a less damaging target, the adversary can respond by shifting to a more damaging target.

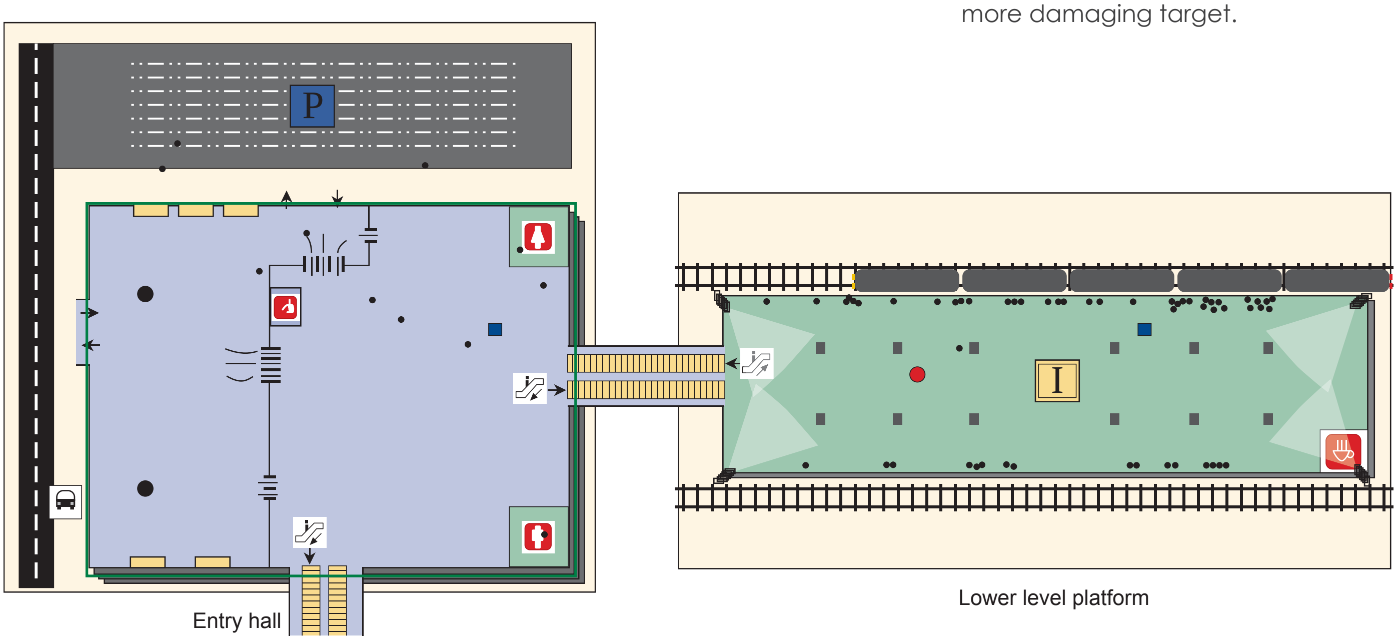

**Figure 2.** The agent-based model. The patrols are blue; the adversary is red; and the pedestrians are black.